Spectral Classification of Ices in the Outer Solar System: Final Report

Eliot Young, *Southwest Research Institute;* Erzsebet Merenyi, Liliy Zhang, *Rice University*; Will Grundy, *Lowell Observatory*

15-FEB-2009

PROJECT GOAL

Get surface properties of ices on distant objects from small details in their spectra. Specifically, we are developing a neural-net spectral classifier to determine:

- temperature
- crystalline structure
- composition
- mixing (e.g., CH_4 in N_2)
- grain size
- scattering phase functions

PROJECT OVERVIEW

We are currently concentrating on spectra of H_2O, CH_4, N_2, and CO ices. These ices are common throughout the outer solar system and can be used to determine many surface properties of icy objects.

We have obtained laboratory optical constants (real and imaginary indices of refraction) for crystalline water, amorphous water ice, methane, methane dissolved in nitrogen ice, isolated nitrogen ice, and CO ice. For most (but not all) of these species we have optical constants over a useful range of temperatures.

We use a Hapke model (a simplified radiative transfer code for light refelected by solid surfaces) to generate synthetic spectra from the optical constants. Grain size, mixing states, and scattering phase functions are free parameters which are inputs to the Hapke code.

We have applied the neural-net classifier to crystalline water ice. The challenge is to separate the effects of grain size from temperature, since both affect the spectra in similar ways. The neural net is composed of two stages: a Self-Organizing Map (SOM) and a Widrow-Hoff back-propagation network.

Most of this project has concentrated on recovering the grain size and temperature of water ice. Previous work on this project helped lead to an award from NASA's Outer Planets Program to study amorphous and crystalline water ice.

INITIAL WORK & RESULTS

In previous years this project has entailed the following tasks:

- Acquisition of optical constants for a series of ices (N_2, CO, CO_2, CH_4, H_2O) over a range of temperatures.
- Generation of model spectra (Fig. 1) for species using a Hapke code (a radiative transfer code to calculate reflectance spectra from icy surfaces).
- Training of a 2-layer neural network using the model spectra (Fig. 2). The first stage is an unsupervised self-organizing map (SOM). The second stage is a conventional back-propagation network.
- Testing the sensitivity of the neural network using "jackknife tests" (Fig. 3)

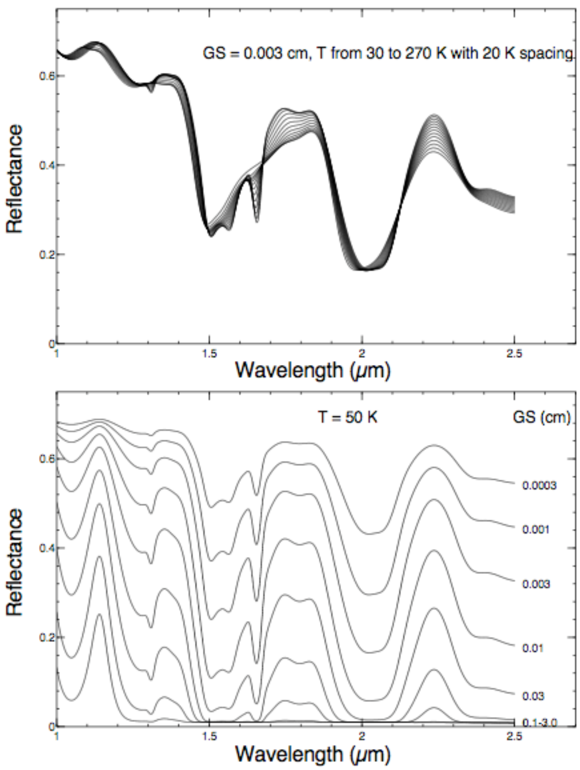

Fig 1: Synthetic crystalline ice spectra used to train the neural net. The top panel shows spectral variations that depend on temperature (for a grain size of 0.003 cm); the bottom panel is a survey of grain sizes at T = 50°C.

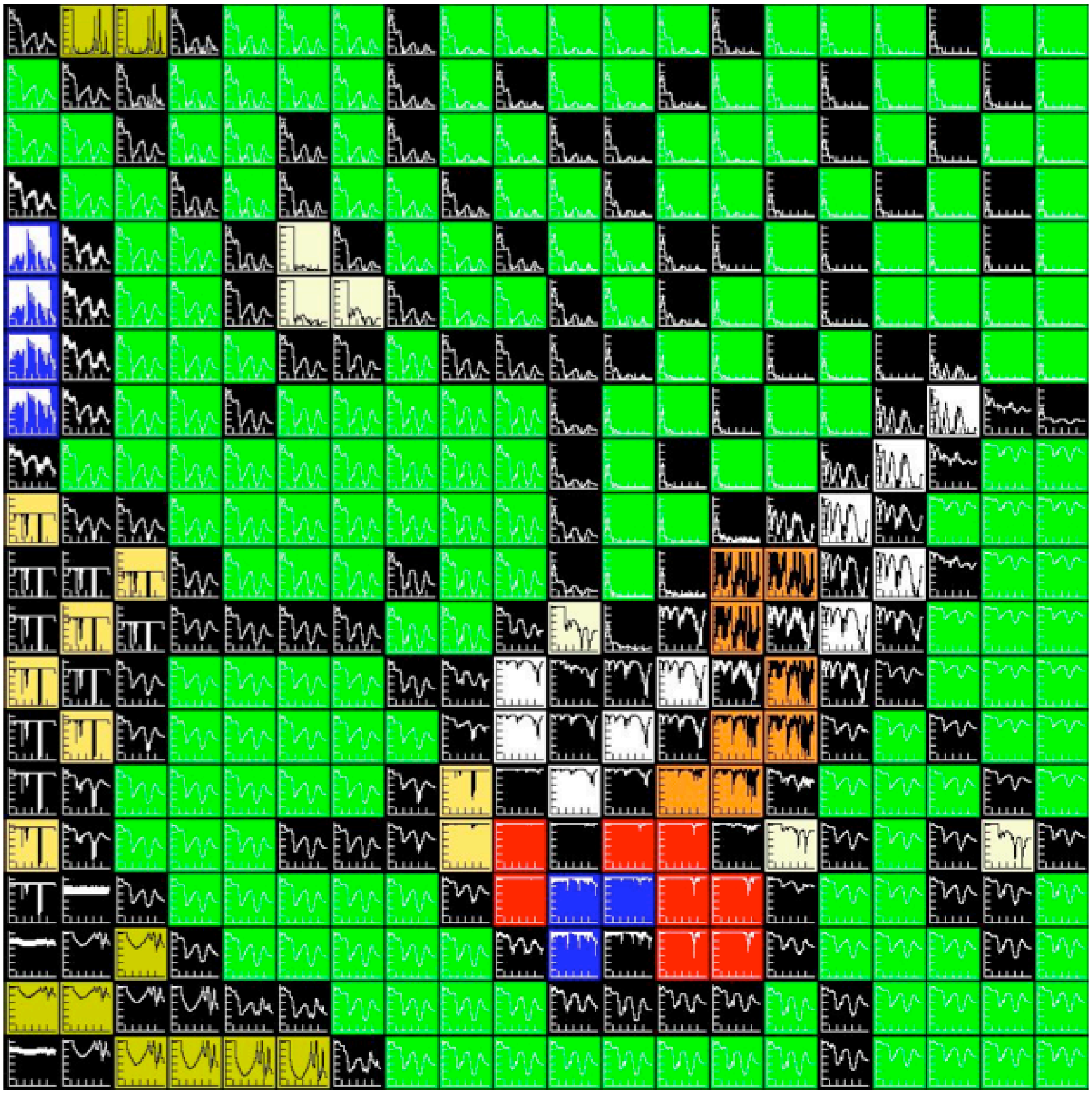

Fig 2: In this self-organizing map, the *green* areas represent crystalline water ice spectra. Nine grain sizes and 126 temperatures (2K spacing) were used to generate the water ice training set. This SOM shows us that the eight compositional candidates (different colors, respresenting spectra of ices of H_2O, CO, CO_2, CH_4, N_2,) are easily distinguished, and that within the green areas, grain size characterization dominates over temperature.

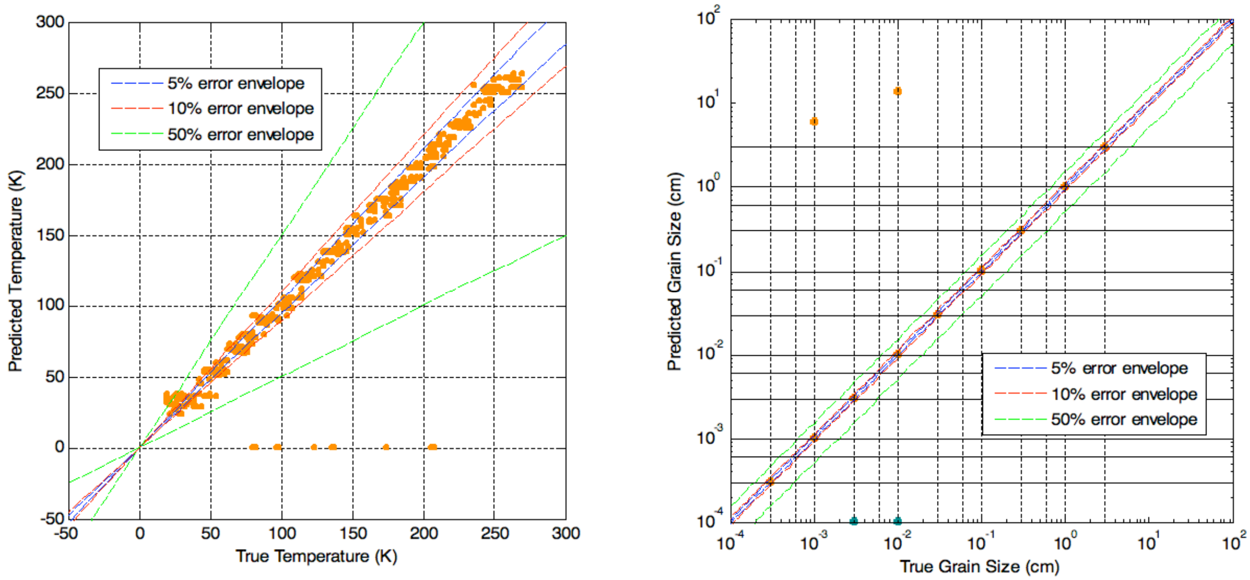

Fig 3: The ability to classify temperature and grain size based on the results of "jackknife" tests (Use a fraction of the training set (2/3) to train the neural net, then test the neural net with the remaining 1/3 of the spectra; repeat with different training/test sets.) These initial tests with crystalline water ice showed that the grain sizes were always correctly classified, but temperatures were classified at the ~75% level.

RECENT WORK & RESULTS

Recently we have concentrated on crystalline water ice, a common surface constituent throughout the outer solar system. The spectrum of crystalline water ice depnds on both temperature and grain size, but larger grains and higher temperatures, to first order, both serve to deepen spectral absorption features. One of the main challenges of this project was to determine whether a neural net spectal classifier could reliably separate the two correlated parameters, temperature and grain size.

The SOM-hybrid neural net (Fig. 4) was used to predict the two parameters.

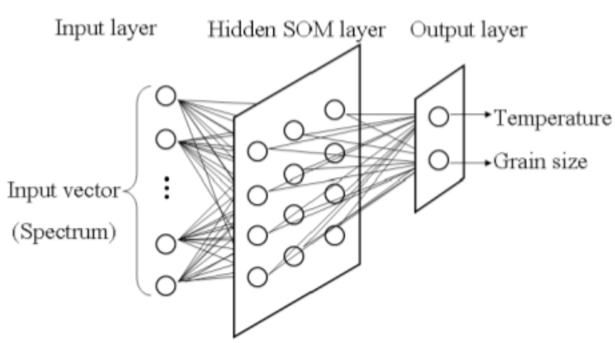

Fig 4: The SOM-hybrid supervised neural architecture.

This neural supervised architecture is a fully connected feed-forward network, with an SOM as the middle layer and an output layer connected to the SOM by the Widrow-Hoff rule. A 230-band spectrum (from 1 - 2.5 μm) is taken as an input vector during every learning step, approximately the range and resolution of infrared image cubes that will be obtained by the New Horizons spacecraft.

In the first stage of learning, i.e. unsupervised learning, the middle SOM layer captures the structure of the data manifold. The knowledge represented on the SOM is then utilized in the second stage of the learning, i.e. supervised learning in the output layer, to achieve a good prediction accuracy for the two physical parameters. The dominance of one parameter, grain size, is demonstarted in the SOM (Fig. 5) by clearly separated clusters organized by grain size. Within these clusters are vectors that span a range of temperatures.

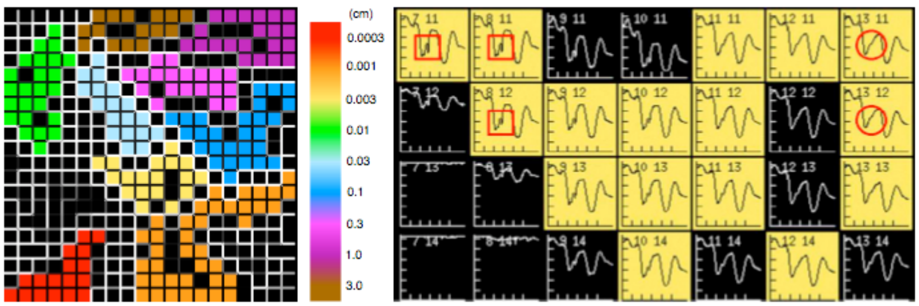

Fig 5: *Left panel*: the 20 x 20 SOM trained with synthetic spectra of crystalline water ice. The superimposed colors represent known grain sizes from the training set. The black cells represent weight vectors that do not represent any training data. *Right panel*: a subframe from the center of the left panel, with yellow vectors representing 0.003 cm grain sizes. In this subgroup there is a continuous change in spectral shapes from left to right caused by increasing temperatures. The red boxes and circles indicate low and high temperatures, respectively. (The number pairs show the grid locations in the SOM.)

The distribution of temperatures within grain size-dominated clusters suggests an alternative to the common "winner take all" (WTA) strategy in which one SOM neuron will fire (have an output signal of 1) in response to a given nput vector, and send a zero otherwise. The other possibility is to divide the winning credit among k SOM nodes by assigning each an output value that is inversely proportional to the distance of its associated prototype vector to the input vector (such that the credits add up to 1). This can be called "interpolation mode". For this study the number of neighbors small; we set $k = 3$.

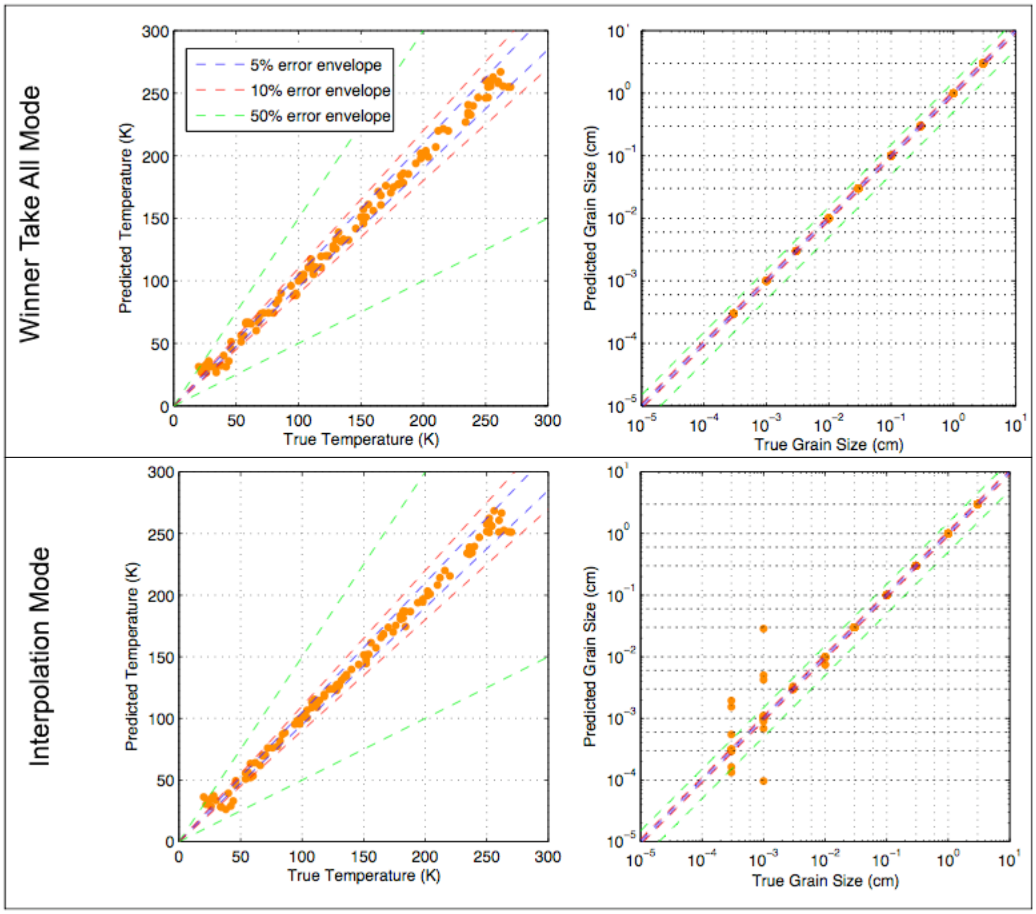

Fig 6: *Top panel*: correlations for True vs. Predicted Temperature (left) and True vs. Predicted Grain Size (right) fir the Winner Take All mode. *Bottom panel*: same as top panel, but the SOM is interrogated in Interpolated Mode. The blue, red and green dashed lines indicate 5%, 10% and 50% error envelopes, respectively. Temperature has a smaller prediction error in *Interpolation Mode*. The prediction of grain size is better in the *Winner Take All Mode*.

The nature of the two physical parameters cause their different dominance on the SOM clustering. As a result, they work against each other in the supervised learning when using the same interpolation mode. Based on the above observations, we can create a "cojoined twin" machine that exploits this duality to improve the overall prediction. It combines the advantages of the WTA and interpolation modes, both "feeding" from the same SOM knowledge.

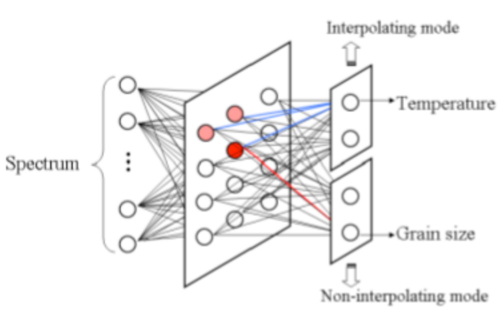

Fig 7: Conceptual diagram of the cojoined twin machines

FUTURE DIRECTIONS

There are a few questions we would like to investigate with this neural net-based spectral classification tool. Like good engineers, we will try to break the tool, specifically, we will assess how the prediction rate degrades in the presence of noisy spectral observations. We will also assess how the prediction rate suffers in the presence of a *range* of grain sizes. This work will be supported under a NASA OPR subcontract to study amorphous and crystalline water ice throughout the solar system.

Under the OPR subcontract, we will feed spectral observations of Ganymede, Callisto, Europa, and several saturnian satellites (from Keck, Cassini, and Galileo observations) to the trained net to help assess surface water ice temperatures and grain sizes.